

IMPLEMENTASI PRINCIPAL COMPONENT ANALYSIS UNTUK IDENTIFIKASI CITRA TANDA TANGAN

Resty Wulanningrum¹⁾, Ema Utami²⁾, Armadyah Amborowati³⁾

¹⁾ Universitas Nusantara PGRI Kediri

^{2), 3)} Magister Teknik Informatika STMIK AMIKOM Yogyakarta

¹⁾ Jl. KH. Achmad Dahlan 76, Kediri, Jawa Timur 64112

^{2), 3)} Jl Ring road Utara, Condongcatur, Sleman, Yogyakarta 55281

Email : resty0601@gmail.com¹⁾, ema.u@amikom.ac.id²⁾, armadyah.a@amikom.ac.id³⁾

Abstrak

Dalam kehidupan sehari-hari, tanda tangan digunakan sebagai identifikasi dari pemilik tanda tangan. Keberadaan tanda tangan dalam sebuah dokumen menyatakan bahwa pihak yang menandatangani, mengetahui dan menyetujui seluruh isi dari dokumen. Sebuah tanda tangan dapat diolah untuk dilakukan identifikasi menggunakan pengolahan citra digital. Salah satunya menggunakan bahasa pemrograman Delphi 7. Selain itu bisa ditambahi dengan metode *Principal component Analysis (PCA)* untuk proses ekstraksi dan *Euclidean Distance* untuk identifikasi dengan mencari jarak terdekat dari citra training dengan citra testing. Penggunaan partisi citra dan nilai threshold akan menghasilkan hasil identifikasi citra tanda tangan yang bervariasi untuk proses analisa yang akan digunakan. Berdasarkan hasil 4 skenario, didapatlah hasil terbaik pada skenario 2 dan 4 dengan nilai threshold sebesar 0-255 menggunakan reduksi dimensi 20 yaitu sebesar 98%.

Kata kunci: PCA, Partisi citra, Euclidean Distance

1. Pendahuluan

Tanda tangan (*signature*) atau Paraf adalah tulisan tangan, yang diberikan gaya tulisan tertentu dari nama seseorang atau tanda identifikasi lainnya yang ditulis pada dokumen sebagai sebuah bukti dari identitas dan kemauan. Tanda tangan berlaku sebagai segel. Fungsi tanda tangan adalah untuk pembuktian. Dalam kehidupan sehari-hari, tanda tangan digunakan sebagai identifikasi dari pemilik tanda tangan. Keberadaan tanda tangan dalam sebuah dokumen menyatakan bahwa pihak yang menandatangani, mengetahui dan menyetujui seluruh isi dari dokumen. Pembubuhan tanda tangan sering dijumpai pada kegiatan administrasi perbankan, seperti: transaksi penarikan uang secara tunai, penyetoran, kliring giro dan transaksi perbankan lainnya. Pada zaman teknologi ini, pencocokan karakteristik tanda tangan dengan pemiliknya dapat dilakukan dengan menggunakan komputer, sehingga akan menghemat waktu bila dibandingkan dengan melakukannya secara manual. Caranya adalah dengan menggunakan *Euclidean*

Distance untuk proses klasifikasi dan *PCA (Principal Component Analysis)* untuk proses ekstraksi.

2. Pembahasan

Pada penelitian yang digunakan menggunakan 2 metode yaitu *PCA* dan *Euclidean Distance*

2.1 Tinjauan Pustaka

Pada literature dari International Journal of Biometrics and Bioinformatics (IJBB), Volume (5) : Issue (4) : 2011, oleh M. H. Sigari, M. R. Pourshahabi & H. R. Pourreza, dengan judul "Offline Handwritten signature Identification and verification using multi-resolution Gabor Wavelet" menghasilkan sistem yang diusulkan memiliki kinerja yang lebih tinggi dalam identifikasi dan verifikasi tanda tangan dengan kebangsaan yang berbeda karena independensi dari bentuk dan struktur tanda tangan. Hal ini diverifikasi dengan menguji sistem yang diusulkan pada 4 database tanda tangan dengan berbagai kebangsaan termasuk Iran (Persia), Afrika Selatan, Turki dan tanda tangan Spanyol. Pada penelitian yang akan dilakukan adalah object yang digunakan untuk penelitian adalah citra tanda tangan dari kebangsaan indonesia, pada suku Jawa yang diambil secara random, khususnya daerah Kediri.

Pada jurnal dengan judul "Pengenalan sidik jari manusia dengan matriks kookurensi aras keabuan (*gray level co-occurrence matrix*)" oleh Reza Syauqi Falasev, Achmad Hidayatno, R. Rizal Isnanto di tahun 2011, dari hasil pengujian dan analisis dapat disimpulkan bahwa tingkat pengenalan tertinggi pada program pengenalan sidik jari dengan menggunakan matriks kookurensi aras keabuan ini sebesar 83,3% yang diperoleh dengan penggunaan sudut jamak ($0^0+45^0+90^0+135^0$) dan jarak 1 piksel dengan citra sidik jari yang tersimpan dalam basisdata sebanyak dua belas citra. Pada penelitian yang akan dilakukan menggunakan nilai ambang keabuan untuk mengetahui intensitas keabuannya.

2.2 PCA dari Sudut Pandang Statistika

Principal Component Analysis merupakan suatu perhitungan standar modern yang digunakan untuk analisis data pada macam-macam field atau multi dimensi sekumpulan data (dataset) khususnya pada bidang komputer grafik, karena kemudahannya dengan

tidak membutuhkan parameter khusus dalam ekstraksi informasi yang berhubungan terhadap sekumpulan data yang meragukan [1]. Dengan usaha minimal PCA menyediakan alur bagaimana mengurangi kumpulan data yang kompleks ke dalam dimensi lebih kecil.

Algoritma ini diperkenalkan pada tahun 1933 oleh H. Hotelling, oleh karena itu sering juga disebut transformasi Hotelling. Untuk pertama kalinya PCA dikembangkan oleh para ahli statistik untuk mengeksplorasi hubungan sejumlah variable kualitatif yang dikembangkan oleh Karl Pearson pada tahun 1901. Karl Pearson membuat makalah yang menganalisis matrik korelasi yang berasal dari pengukuran tujuh variabel fisik untuk tiap orang dari 3000 kriminal [2].

2.3 Metode Euclidean Distance

Metode *Euclidean* yaitu metode klasifikasi tetangga terdekatnya dengan menghitung jarak antara dua buah obyek, metode ini disebut juga jarak *Euclidean* [3]. Rumus penghitungan jarak ditulis sebagai berikut:

$$d_e = \sqrt{\sum_{k=1}^m (fd_{i,k} - k_j)^2}$$

Keterangan :

- d_e : jarak euclidean
- fd_i : bobot citra pelatihan
- k_j : data bobot test
- m : jumlah data pelatihan

2.4 Gambaran umum Obyek Penelitian

Data citra yang digunakan dalam penelitian ini adalah data tanda tangan yang diambil dari 5 orang usia antara 27-55 tahun menggunakan sebuah kertas yang diisi tanda tangan kemudian dilakukan scanner. Dari hasil scanner akan dilakukan pemotongan gambar untuk setiap satu data citra tanda tangan berukuran 100 piksel x 100 piksel. Pada masing- masing orang diambil sebanyak 30 sampel. Data citra disimpan dalam *file* dengan format *bitmap*.



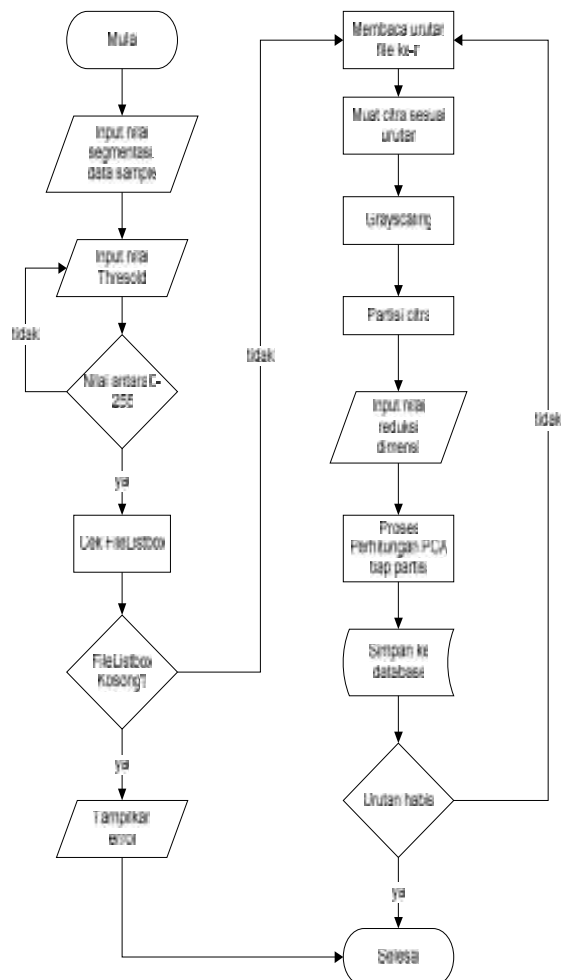
Gambar 1. Citra tanda tangan

2.5 Analisa dan Rancangan sistem

Penelitian ini membutuhkan suatu aplikasi atau perangkat lunak untuk menunjang hasil dari penelitian

ini. Sistem yang dirancang memiliki beberapa fungsi, yaitu untuk proses akuisisi data, pengolahan citra untuk data *training*, pengolahan citra untuk data *testing* dan menyajikan hasil identifikasi.

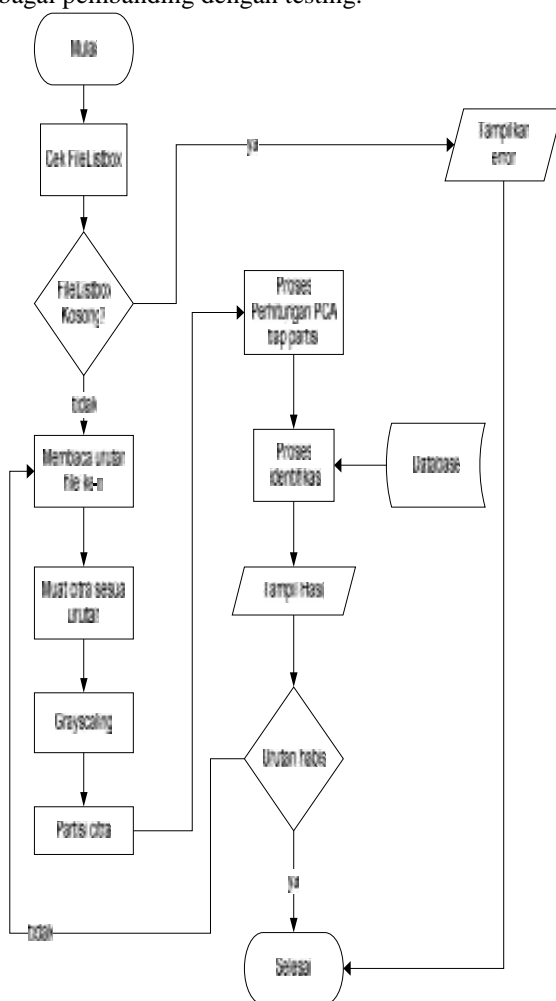
Pada pengolahan citra untuk data *training* dan *testing* memiliki kesamaan proses. Namun pada proses data *training* hasilnya disimpan dalam database. Database yang digunakan pada aplikasi ini adalah jenis database yang menggunakan format.tbl. File ini hanya digunakan untuk menyimpan data nilai matriknya saja. Alur program untuk proses *training* disajikan pada gambar 2 dan proses *testing* pada gambar 3.



Gambar 2. Alur data *training*

Alur dari gambar 2 adalah, pertama kita input nilai segmentasi atau partisi citra serta jumlah sampel tiap kelas, kemudian memasukkan nilai thresold dari 0-255. Sistem akan membaca citra dari filelistbox yang telah disediakan. Jika filelistbox kosong, maka akan muncul eror kemudian selesai. Jika filelistbox tidak kosong, maka dia akan membaca urutan 1 sampai ke-n kemudian dilakukan proses *grayscale* dilanjutkan dengan proses partisi citra. Setelah pembacaan citra pada filelistbox selesai, input nilai reduksi dimensi yang akan digunakan. Setelah proses reduksi dimensi baru dilakukan proses PCA di tiap partisi. Simpan nilai pada

database sebagai nilai training yang akan digunakan sebagai pembanding dengan testing.

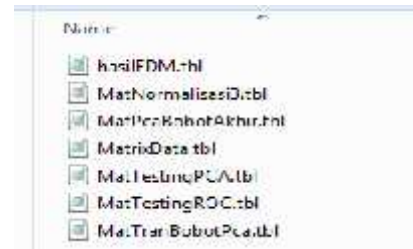


Gambar 3 Alur data Testing

Pada gambar 3 alur data *testing*, citra yang akan dilakukan proses *testing* atau ujicoba akan ditampilkan pada filelistbox. Image previewnya akan ditampilkan pada TImage. Pada saat *testing*, citra yang dipilih akan dilakukan proses *grayscaleing* terlebih dahulu, baru dipartisi selanjutnya proses PCA (*principal Component Analysis*). Setelah itu akan dilakukan identifikasi dari dengan membandingkan nilai citra *testing* dengan citra *training* yang sudah disimpan nilai matriknya. Hasil identifikasinya keluar pada message box.

2.6 Analisa hasil

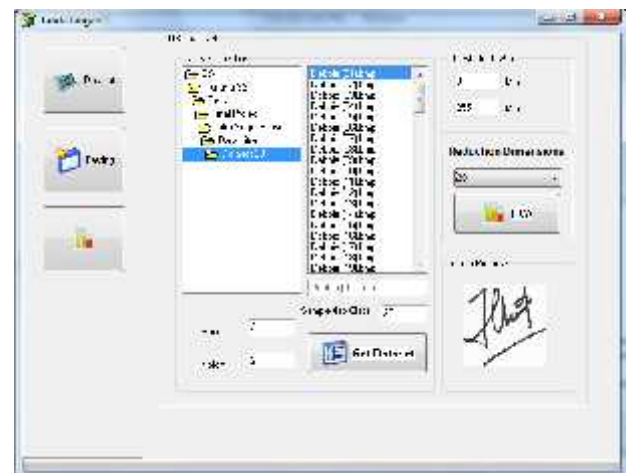
Berdasarkan hasil uji coba implementasi identifikasi tanda tangan menggunakan bahasa pemrograman Delphi 7 dengan SDL Component 7.2. Sistem identifikasi tangan tangan ini diimplementasikan dengan menggunakan sistem operasi Windows 7 Professional Service Pack 1. Untuk menyimpan hasil nilai matrik menggunakan file .tbl. File .tbl (file dengan ekstensi .tbl) adalah file teks yang berisi nilai matrik pada citra *training*. Seperti ditunjukkan oleh gambar 3.



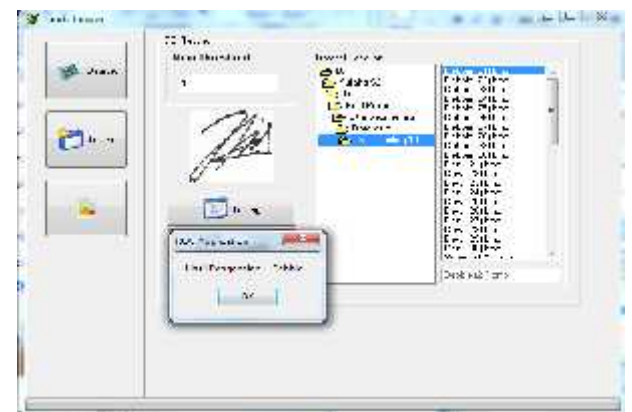
Gambar 4. File.tbl yang berisi nilai matrik citra *training*

Pada saat pertama aplikasi dijalankan akan mengambil data bobot pada *mat.PcaBoobotAkhir.tbl*. File ini akan menyimpan data matrik untuk yang terakhir kali pada proses *training*.

Hasil interface ditunjukkan oleh gambar 5 untuk proses *training* dan gambar 6 untuk proses *testing*.



Gambar 5. Form Training



Gambar 6. Form Testing

Dalam pengujian sistem identifikasi citra tanda tangan menggunakan skenario 1 ujicoba yang dilakukan pada data *training* sebanyak 10 data citra dengan ukuran 100 piksel x 100 piksel dengan menggunakan partisi 2x2 dan data *testing* sebanyak 20 data diluar data *training* ditunjukkan oleh tabel 1. Hasil akurasi terbaik adalah pada nilai thresold 0-255 dengan reduksi dimensi sebesar 20, yaitu hasil akurasinya 97%. Semakin besar reduksi dimensi, maka akan semakin rendah akurasinya.

Tabel 1. Skenario 1

Threshold	Data Training	Data Testing	Reduksi Dimensi	Akurasi %
47- 128	10	20	20	76
47- 128	10	20	25	72
113- 219	10	20	20	74
113- 219	10	20	25	70
0-255	10	20	20	97
0-255	10	20	25	96

Pada skenario ujicoba 2, menggunakan data *training* sebanyak 10 data citra menggunakan partisi 5x5 dan data citra berukuran 100 piksel x 100 piksel dengan data *testing* 20 data diluar data *training*. Hasil ujicoba ditunjukkan pada tabel 2. Hasil akurasi terbaik adalah pada nilai threshold 0-255 dengan reduksi dimensi sebesar 20, yaitu hasil akurasinya sebesar 98%.

Tabel 2. Skenario 2

Threshold	Data Training	Data Testing	Reduksi Dimensi	Akurasi %
47- 128	10	20	20	77
47- 128	10	20	25	72
113- 219	10	20	20	76
113- 219	10	20	25	71
0-255	10	20	20	98
0-255	10	20	25	97

Pada skenario ujicoba 3, menggunakan data *training* sebanyak 20 data citra menggunakan partisi 2x2 dan data citra berukuran 100 piksel x 100 piksel dengan data *testing* sebanyak 10 data diluar data *training*. Hasil ujicoba ditunjukkan pada tabel 3. Hasil akurasi terbaik adalah pada nilai threshold 0-255 dengan reduksi dimensi sebesar 20, yaitu hasil akurasinya sebesar 94%. Nilai akurasinya lebih rendah dibanding pada skenario 1, dikarenakan data *testing* yang digunakan lebih besar dibanding data *training*. Sehingga berpengaruh pada hasil akurasinya.

Tabel 3. Skenario 3

Threshold	Data Training	Data Testing	Reduksi Dimensi	Akurasi %
47- 128	20	10	20	86
47- 128	20	10	25	82
113- 219	20	10	20	84
113- 219	20	10	25	80
0-255	20	10	20	94
0-255	20	10	25	90

Pada skenario ujicoba 4, menggunakan data *training* sebanyak 20 data citra menggunakan partisi 5x5 dan data citra berukuran 100 piksel x 100 piksel dengan data *testing* sebanyak 10 data diluar data *training*, hasil ujicoba ditunjukkan pada tabel 4. Hasil akurasi terbaik adalah pada nilai threshold 0-255 dengan reduksi dimensi sebesar 20, yaitu hasil akurasinya sebesar 98%.

Tabel 4. Skenario 4

Threshold	Data Training	Data Testing	Reduksi Dimensi	Akurasi %
47- 128	20	10	20	86
47- 128	20	10	25	82
113- 219	20	10	20	86
113- 219	20	10	25	82

0-255	20	10	20	98
0-255	20	10	25	97

Berdasarkan hasil 4 skenario, didapatkan hasil terbaik pada skenario 2 dan 4 dengan nilai threshold sebesar 0-255 menggunakan reduksi dimensi 20 yaitu sebesar 98%.

3. Kesimpulan

Berdasarkan data dan uraian yang disajikan maka dapat disimpulkan sebagai berikut:

1. Sistem identifikasi tanda tangan menggunakan Principal Component Analysis (PCA) dan jarak Euclidean Distance dapat berfungsi dengan baik.
2. Pemberian nilai threshold untuk sistem identifikasi citra tanda tangan mempengaruhi nilai akurasi pada proses pengenalan citra tanda tangan. Akurasi terbaik pada skenario 2 dan 4 dengan nilai threshold sebesar 0-255 menggunakan reduksi dimensi 20 yaitu sebesar 98%..
3. PCA merupakan salah satu metode yang dapat digunakan untuk mengkonversi dan menormalisasi suatu citra, sehingga citra tersebut dapat dilakukan proses ekstraksi untuk identifikasi tanda tangan.
4. Semakin tinggi partisi atau *region base* pada citra juga berpengaruh terhadap akurasi. Semakin besar pembagian partisi, maka akan semakin tinggi pula akurasinya, karena informasi yang dimiliki untuk proses PCA akan semakin detail, sehingga hasilnya lebih bagus.

Daftar Pustaka

- [1] Smith, Lindsay. "A Tutorial on Principal Component Analysis". 2002. Sumber http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf. Diakses pada 2013.
- [2] Leon, Steven J.. " Aljabar Linear dan Aplikasinya". Erlangga, Jakarta. 2001.
- [3] Navarette, Pablo dan Javier Ruiz- del- Solar."Self-Organizing Maps-Kohonen",2003.
- [4] Falasev, Reza Syauqi, Achmad Hidayanto, R.Rizal Isnanto. "Pengenalan sidik jari manusia dengan menggunakan Matrik Kookurensi Aras Keabuan (Gray Level Co-Occurecy Matrix)". Seminar nasional Aplikasi Teknologi Informasi (SNATI). 2011.
- [5] Sigari.M.H.M.R.Pourshahabi & H.R.Pourreza. "Offline handwritten Signature Identification and Verification Using Multi Resolution Gabor Wavelet". International Journal of Biometrics and Bioinformatics (IJBB), Volume(5):Issue(4),234-248.2011

Biodata Penulis

Resty Wulanningrum, memperoleh gelar Sarjana Komputer (S.Kom), Jurusan Teknik Informatika Universitas Trunojoyo Madura, lulus pada tahun 2011. Saat ini sebagai Staf Pengajar Program Sarjana Teknik Informatika UNP Kediri.

Emu Utami, memperoleh gelar Sarjana Sains (S.Si) dari Ilmu Komputer Universitas Gadjah Mada Yogyakarta pada tahun 1997, memperoleh gelar Magister Komputer (M.Kom) dari Program Pascasarjana Ilmu Komputer Universitas Gadjah Mada Yogyakarta pada tahun 2002,

memperoleh gelar Doktor (Dr) dari Program Doktorat Ilmu Komputer Universitas Gadjah Mada Yogyakarta pada tahun 2010. Saat ini menjadi Wakil Direktur I Bidang Akademik Program Pascasarjana STMIK AMIKOM Yogyakarta.

Armadyah Amborowati, memperoleh gelar Sarjana Komputer (S.Kom), Jurusan Sistem Informasi STMIK AMIKOM Yogyakarta. Memperoleh gelar Magister Engineering (M.Eng) Program Pasca Sarjana Magister Teknologi Informasi, Teknik Elektro, Universitas Gajah Mada Yogyakarta. Saat ini menjadi Dosen di STMIK AMIKOM Yogyakarta.

